

Available online at www.sciencedirect.com**ScienceDirect**

Procedia Technology 25 (2016) 411 – 419

Procedia
 Technology

Global Colloquium in Recent Advancement and Effectual Researches in Engineering, Science and Technology (RAEREST 2016)

BF-ASIFT-2DPCA and ABF-ASIFT-2DPCA for Face Recognition

Vinay. A¹, Joella Regina Mathias², Anita M. Fathepur², K N Balasubrmanaya Murthy¹,
S Natarajan²

¹Department of Computer Science, PES University, Bengaluru, Karnataka

²Department of Information Science, PES Institute of Technology, Bengaluru, Karnataka

Abstract

Walk into a shopping mall, the clerk will know your name and other details. The mall's cameras will know your identity by indexing you up in both their gallery and a wide gamut of marketing database they you may have subscribed to. Customers profile information including name, address, sale interests are displayed to the clerk. Further the clerk can infer what sort of sales pitches you're most vulnerable to and how profitable a customer you are to that shopping mall. Similarly, walk by a policeman, they will identify your name, address and criminal record, if any. The underlying technology here is automated face recognition. This paper proposes two variations of the Affine-SIFT technique for face recognition. It outlines the applications of two filters namely the Bilateral filter and the Adaptive Bilateral filter. The quality of the image is enhanced before performing feature extraction and matching, so that matching will be more accurate. The use of filters not only increases the image quality, but also leads to more accurate key point matches which are crucial for image matching. The filters aid in pre-processing the image before feature extraction by smoothening the image, while preserving the edges and resulting in sharper images. This paper provides a comparative study of the various filters that can be used for pre-processing, some resulting in complete blurring of the image, some preserve the edges and some sharpen the image resulting in improved matches. It also explains how these pre-processing techniques impact the key point descriptors. The methodology also implements the Two Dimensional Principal Component Analysis (2DPCA) technique for dimensionality reduction of key points. The proposed model shows invariance to different levels of scale, illumination and poses to some extent.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the organizing committee of RAEREST 2016

Keywords: ASIFT; BF- ASIFT-2DPCA, ABF-ASIFT-2DPCA, ASIFT-2DPCA

1. INTRODUCTION

Image processing has been an important part of research since a couple of decades. It continues to find its way in

almost every area in the recent years, be it video, multimedia, security, surveillance, medical and other biological purposes. Face recognition plays a significant role in the image processing area. Facial recognition systems have been developed for identifying and verifying individuals from an image or video frame. This is done using feature matching with respect to a facial database. Facial recognition systems face a wide variety of challenges which include changes in illumination, which correspond to whether the image was taken during day/night, changes in pose of the person, changes in scale, changes in facial expressions and other issues like, whether the person is wearing spectacles, with/without moustache etc.

This paper proposes two variants of the ASIFT (Affine-Scale Invariant Feature Transform) technique [1], mainly the BF-ASIFT (Bilateral Filter ASIFT) and the ABF-ASIFT (Adaptive Bilateral ASIFT) techniques together with PCA (Principal Component Analysis) [3] for dimensionality reduction. The application of these filters before performing feature matching shows substantial improvement in the matching accuracy compared to the ASIFT technique. These filters improve the image quality by avoiding the blurring of edges and boundaries in the image and aid in sharpening of the image which in turn gives better feature matches.

2. RELATED WORK

Since the time the SIFT technique [2] was proposed by David G. Lowe in 1999, a lot of research has been carried out to improve the performance of the proposed model. The original model that was proposed by Lowe consisted of mainly four steps which include detection of scale-space extrema, keypoint localization, orientation assignment and keypoint descriptor. Scale-space extrema detection involves detection of keypoints for all possible levels of scale. These keypoints are then stabilized by localization to sub-pixel accuracy. Eventually, unstable keypoints are eliminated. In the orientation assignment step, dominant orientations for each keypoint based on its local image patch are identified. Finally, a local image descriptor for each keypoint is built, based upon the 4x4 image gradients in its local neighborhood. It is found that the SIFT technique is invariant to image scaling and rotation and partially invariant to illumination and viewpoint changes. A number of SIFT descriptor variants such as PCA-SIFT [4], CSIFT [5], GSIFT [6] and SURF [7] have been developed since then. These methodologies claim to be more robust and distinctive with scaled-down complexity levels.

The PCA-SIFT technique was introduced by Y. Ke et. al. [4] which follows the same steps as in the SIFT technique along with using PCA for dimensionality reduction. The proposed methodology applies PCA to the normalized gradient patch instead of using SIFT's smoothed weighted histograms. Their experiments demonstrate that the PCA based local descriptors in comparison to the standard SIFT representation are more distinctive, more invariant to image deformations, and more compact. In Comparison to the standard SIFT representation, PCA-SIFT is found to be more distinctive and more compact which leads to significant improvements in matching accuracy for both controlled and real-world conditions.

The drawback of SIFT that it is designed mainly for gray images, led to the research work by Alaa E. Abdel-Hakim et al [5] for designing a model for colored images as well. They address the problem of gray images by proposing a colored local invariant feature descriptor by embedding color information in the descriptor and providing robustness with respect to color variations. Instead of representing the input image in the gray space, their approach builds the SIFT descriptors in a color invariant space. With respect to color and photometrical variations, the Colored SIFT (CSIFT) model is more robust than the conventional SIFT.

Mortensen, E.N et. al. [6] proposed a model that adds a global texture vector to the basis of SIFT. The SIFT model fails to consider global context to resolve ambiguities that can occur locally when an image has multiple similar regions and can only find matches between features with unique local neighborhoods. The model suggested in their work presents a feature descriptor that integrates SIFT with a global context vector that adds curvilinear shape information from a much larger neighborhood. Thus it helps reducing mismatches when multiple local descriptors are similar.

3. PROPOSED SYSTEM

1.1. Preprocessing using Bilateral and Adaptive Bilateral filters

Filtering is very necessary method, when it comes to face recognition. We first implemented the Gaussian filter. Gaussian worked fairly well, but carried many disadvantages. It considers the neighbourhood around a pixel and computes its Gaussian weighted average. The Gaussian filter fails to compare pixels with similar intensities and ends up in blurring the edges also. To remove these failures of the Gaussian method, we came up with a very effective solution that is Bilateral Filtering. Bilateral filtering in simple words just means Edge preserving method.

Bilateral filtering dates back to 1995 with the work of Aurich and Weule on nonlinear Gaussian filters. It was later rediscovered by Smith and Brady. The bilateral filtering is non-linear technique that can blur an image, keeping the edges strong. It's a highly effective in noise removal. Gaussian function of intensity difference make sure only those pixels with similar intensity to central pixel is considered for blurring, which in turn helps in preserving the edges since pixels at edges will have large intensity variation. It has ability to decompose an image into different scales without causing any unnecessary outliers.

The use of bilateral filtering has grown rapidly, and this method or technique is very crucial in face recognition. The bilateral filter has several qualities that explain its success. Each pixel is replaced by a weighted average of its neighbors. This makes it easy to implement in any other applications. It depends only on two parameters, the size and contrast of the features. It can be used in a non-iterative manner. Bilateral filtering (BF) smoothens an image whilst preserving the strong edges. It has been used in applications ranging from image de-noising to edge enhancement, exposure correction, tone mapping.

Bilateral filter can be formulated as,

$$I^{Filtered}(x) = \frac{1}{w_p} \sum_{x_i \in \Omega} f_r(\|I(x_i) - I(x)\|) g_s(x_i - x) \quad (1)$$

Where normalization is given by

$$w_p = \sum_{x_i \in \Omega} f_r(\|I(x_i) - I(x)\|) g_s(x_i - x) \quad (2)$$

- I (filtered) is the filtered image;
- I is the original input image to be filtered;
- X are the coordinates of the current pixel to be filtered;
- Ω is the window centered in X ;
- f_r is the range kernel for smoothing differences in intensities. This function can be a Gaussian function;
- g_s is the spatial kernel for smoothing differences in coordinates. This function can be a Gaussian function.

We have seen the bilateral filter, preserve the edge and filters. Further we implemented another method called Adaptive Bilateral Filter (ABF) for sharpness enhancement and noise removal. ABF sharpens an image by increasing the slope of the edges. Compared with the bilateral filter, ABF restored images are significantly sharper. ABF does not involve detection of edge. It is able to smooth the noise and enhance the edges and texture in the image.

The output of the filter in position x , $d(x)$ is given by:

$$d(x) = \frac{\sum_{y \in N(x)} e^{-\frac{\|y-x\|^2}{2\sigma_d^2}} - e^{-\frac{\|t(y)-t(x)\|^2}{2\sigma_r^2}} t(y)}{\sum_{y \in N(x)} e^{-\frac{\|y-x\|^2}{2\sigma_d^2}} - e^{-\frac{\|t(y)-t(x)\|^2}{2\sigma_r^2}}} \quad (3)$$

where $t(x)$ is the noisy image, $N(x)$ is the neighbourhood of x , and σ_d and σ_r are the filter parameters. Improvements to basic BF include Adaptive Bilateral. ABF is obtained using different values of the filter parameters for each pixel, i.e., by defining two maps $\sigma_d = \sigma_d(x, y)$ and $\sigma_r = \sigma_r(x, y)$.

1.2. Feature Extraction using Affine SIFT

During the study of SIFT, we see that it was not taking care of the difference in angle measure of person's face. So we implement ASIFT to yield better results, with different face angles and measures. 'A' in ASIFT stands for Affine invariance. A physical object will have a smooth boundary; its images obtained by cameras in varying positions undergo smooth apparent deformations. We have deployed as it takes care of the deformation, which is the geometric curve. The geometry remains consistent when the image is rotated, scaled or translated. SIFT is invariant to four out of six parameters of affine transform. ASIFT simulates the two camera axis parameters, namely the latitude and the longitude angles. It then applies SIFT which simulates the scale and normalizes the rotation and the translation. The transition tilt which measures the degree of viewpoint change from one view to another is a crucial parameter for evaluating the performance of affine recognition. For example a person's facial image might be captured from different angles. It's a very necessary task to understand these parameters in depth, and apply the same in our face recognition process. The attainable transition tilt is measured for each affine image comparison method.

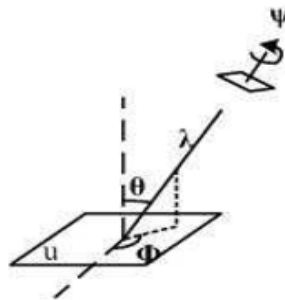


Fig.1. Affine camera model (Image courtesy: [1])

The above diagram helps us to understand the different parameters used in affine camera model. Here we briefly describe each one of them. In the image u is a flat physical object. The parallelogram on the top-right represents a camera looking at u . The angles ϕ and θ are respectively the camera optical axis longitude and latitude. The angle ψ parameterizes the camera spin. λ is a zoom parameter. Absolute tilt measures the tilt between the frontal view and a slanted view. Transition tilt quantifies the amount of tilt between two slanted views. A camera at a finite distance looking at a smooth object is equivalent to multiple local cameras at infinity.

We brief, little more about the camera model. A camera motion of the affine decomposition, ϕ and $\theta = \arccos 1/t$ are the viewpoint angles, ψ parameterizes the camera spin. The camera is assumed to stay far from the image and starts from a frontal view u , i.e., $\lambda = 1$, $t = 1$, $\phi = \psi = 0$. Then the camera is moved parallel to the object's plane, this creates the translation T along the camera axis. The axis meets the plane at the fixed point. It makes an angle ϕ with a fixed vertical plane, called longitude. It then makes an angle θ with the normal to the plane, called latitude. The camera can rotate around its optical axis, making rotational parameter ψ . Last but not least, the camera can move forward or backward, as measured by the zoom parameter λ . The transition tilt is designed to quantify the amount of tilt between two such images.

1.3. Dimensionality Reduction using 2DPCA

Principal component analysis (PCA) was invented in 1901 by Karl Pearson. PCA is mainly used to remove redundancy. It is a variable reduction procedure. It will reduce the variable in to smaller number of variables. PCA aims at reducing the dimensionality, yet tries to retain as much variation as possible in the original data set. This might lead to information loss. Typically, PCA works on 1-dimensional vectors which causes intrinsic problems when employed with high dimensional vector space like face images. On other hand 2DPCA directly works on original face image without transforming into 1D vector.

2D-PCA is popular Dimensionality Reduction and De-noising technique used to represent the key-point features (feature vector extraction). It operates by cutting down the redundancy i.e. reducing the aggregate number that is needed to conclusively announce a match. A suitable value for n (dimensionality) was chosen as 20 because as it's been shown to render favourable results in many literatures.

Let us suppose that X represents an n -dimensional unitary column vector and we perform the projection of an image A ($m \times n$ random matrix) onto X by employing the linear transformation

$$Y = AX \quad (4)$$

This yields an m -dimensional projected vector Y , which represents the projected feature vector of image A .

To ensure accurate projection of X , the total scatter of the projected samples is computed using the trace of the covariance matrix of the projected feature vectors as follows:

$$J(X) = \text{tr}(S_x) \quad (5)$$

where S_x represents the covariance matrix of the projected feature and $\text{tr}(S_x)$ denotes its trace. We require a projection direction X onto which all the samples can be projected so as to maximize the total scatter of the resulting projected samples.

Then the covariance matrix S_x is as follows:

$$S_x = E(Y - EY)(Y - EY)^T = E[AX - E(AX)][AX - E(AX)]^T \\ = E[(A - EA)X][(A - EA)X]^T$$

$$\text{Which yields, } \text{tr}(S_x) = X^T [E(A - EA)^T (A - EA)] X \quad (6)$$

$$\text{We then have } G_T = E[(A - EA)^T (A - EA)] \quad (7)$$

where G_T is the image covariance matrix.

Subsequently the optimal projection vectors of 2DPCA X_1, X_2, \dots, X_d are utilized to perform feature extraction. Let us suppose that for a given image sample A ,

$$Y_k = AX_k \text{ where } k=1, 2, \dots, d. \quad (8)$$

We now obtain a family of projected feature vectors, Y_1, \dots, Y_d (principal component (vectors) of the image A). In this context, the principal component will be a vector for 2DPCA, whereas for conventional PCA it is a scalar. The eigenvectors that correspond to the 20 largest eigen values (20 was chosen as a suitable value as it showed promise in literature) X_1, X_2, \dots, X_{10} are employed as projection axes. The principal component vectors obtained here are utilized to form an $m \times d$ matrix $B = [Y_1, \dots, Y_d]$, which is known as the feature matrix of the image sample A .

1.4. Descriptor Matching using FLANN

After performing Dimensionality Reduction, the descriptors from the query and gallery images are matched using the matching algorithm FLANN [11] to determine whether there is a match or not. Let us suppose that P and G represent the Probe and Gallery images respectively, then $P_k = \{P_{k1}, P_{k2}, P_{k3}, \dots, P_{kn}\}$ where P_k denotes the set of key-points of Probe Images (P) and similarly, $G_k = \{G_{k1}, G_{k2}, G_{k3}, \dots, G_{kn}\}$ represents the set of key-points of Gallery images (G). Further suppose that $PD = \{PD_1, PD_2, PD_3, \dots, PD_{mn}\}$ represents a set of descriptors of image P , while $GD = \{GD_1, GD_2, GD_3, \dots, GD_{mn}\}$ denotes a set of descriptors of image G . We then have, $MD = \{MD_{11},$

MD12, MD13....MDmn} which is the set of matched descriptors, as represented in Table 1 and Table 2. We compare the feature vectors extracted from the query and gallery faces to verify whether there is a match by applying the FLANN [8].

4. EXPERIMENTAL SETUP

Our experiments are carried out using the ORL database and the LFW database of faces which are described below:

1.5. The ORL database

Known as the Database of Faces, the ORL database was developed at the Olivetti Research Laboratory. This database consists of 40 subjects with 10 different images of them arranged in separate folders. The images were taken at different times, varying illumination conditions, changes in facial expressions (open/closed eyes, smiling/non-smiling) and other facial details (glasses/no-glasses). All the images are taken against a dark homogeneous background and the subjects are in up-right, frontal position (with tolerance for some side movement).



Fig.2.The ORL database

1.6. The LFW database:

LFW (Labelled Faces in the Wild) is a database of face photographs designed for studying the problem of unconstrained face recognition. It contains more than 13,000 images of faces collected from the web. Each face has been labelled with the name of the person pictured. 1680 of the people pictured have two or more distinct photos in the data set. In comparison with the ORL database, LFW database contains more real-time images taken at varying conditions and not just a dark homogeneous background.



Fig.3.The LFW database

Experiments are carried out for the methodologies mentioned and for the above databases. We started with the

standard ASIFT technique with Gaussian blurs, which smoothens the image and ends up blurring the edges as well. To overcome this drawback, other filters are implemented. The bilateral filter is used next for the experiment which preserves the edges, followed by Adaptive Bilateral filter which results in sharper images. Principal Component Analysis is used for dimensionality reduction to overcome the high dimensionality of the image space.

5. RESULTS AND DISCUSSION

The table 1 and 2 below shows the matches and mismatches for the techniques discussed in the paper. We see that the application of filters and dimensionality reduction techniques shows substantial improvement in the keypoint matches.

Table 1. Table showing matches for ORL database













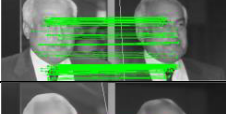

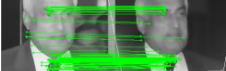
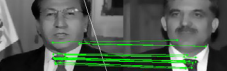
Technique used	Match	Mismatch
ASIFT		
ASIFT-PCA		
BF-ASIFT-PCA		
ABF-ASIFT-PCA		

Table 2. Table showing matches for LFW database

Technique used	Match	Mismatch
ASIFT		
ASIFT-PCA		
BF-ASIFT-PCA		
ABF-ASIFT-PCA		

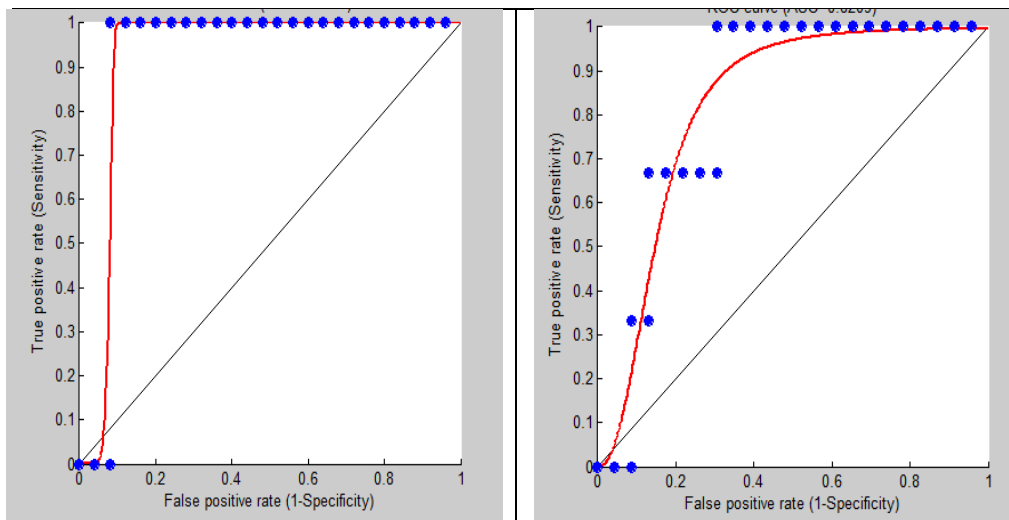
Various metrics used for evaluation of the matching techniques are given below. These values are calculated based on the true positive, false positive, true negative and false negative matches obtained.

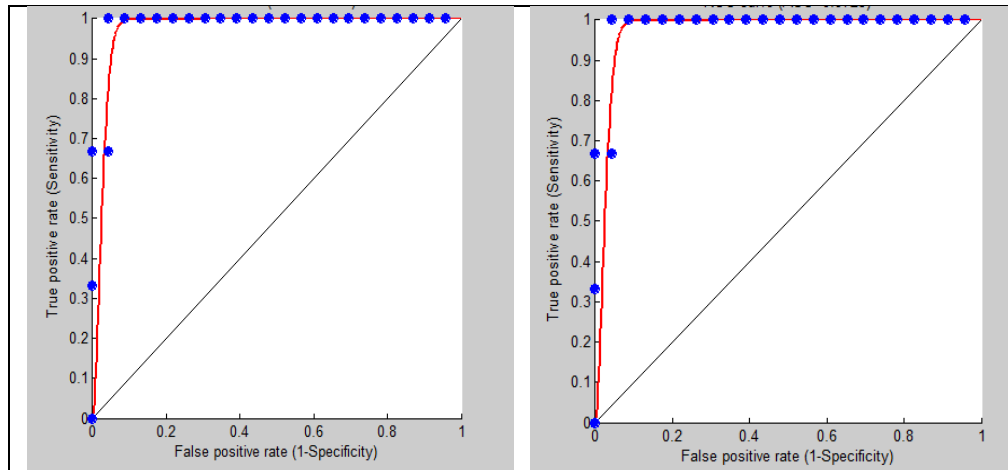
Table 3. Tabulated results

Technique	Accuracy= (TP+TN)/Total	Precision= TP/(TP+FP)	Recall= TP/(TP+FN)	F1 Score= 2* ((Precision*Recall) / (Precision+Recall))
ASIFT	93.75%	50%	92%	0.647
ASIFT-2DPCA	87.3%	76.47%	46.42%	0.577
BF-ASIFT-2DPCA	97.5%	92.68%	84.44%	0.883
ABF-ASIFT-2DPCA	97%	86.54%	90%	0.882

The 2DPCA technique removes the un-modelled distortions by discarding the components/eigen vectors with lower eigen values. But discarding the lower components may result in information loss to some extent. The Bilateral filter results in better accuracy compared to the ASIFT with the Gaussian blurring because it considers relative intensity of the pixels before blurring. So the edges are preserved when applying the Bilateral filter. The Adaptive Bilateral filter makes the images sharper by increasing the slope of the edges. High precision value in case of BF-ASIFT-2DPCA and ABF-ASIFT-2DPCA shows that these algorithms returned substantially more relevant results than irrelevant, while high recall shows that these algorithms returned most of the relevant results. The F1 score conveys the balance between the precision and the recall.

Table 4. ROC curve to analyze performance





A Receiver Operating Characteristic (ROC) curve is a broad way of analyzing the performance of a face recognition system. This curve depicts the confidence of false match rate with the false non-match rate as the system threshold on match score is changed. Table 4 shows the ROC curves corresponding to various methods with respect to the ORL dataset.

6. Conclusion and Future work

This paper proposed two approaches for face recognition based on bilateral filter, ASIFT and 2DPCA. The suggested methods offer invariance to scale, illumination, expression and pose. However, there are many directions for further research in deriving invariant and distinctive image features. Systematic testing is needed on data sets with full 3D viewpoint and illumination changes. Experiments show that Bilateral ASIFT 2DPCA gives the best results in terms of accuracy and precision when tested on the ORL database. The suggested methodology can be extended to face-tagging in group images. This involves identify feature matches of a person in a group image based on the training data. Research can be carried out in this regard and can be implemented in various other applications of face recognition.

References

1. Morel, J.M., Yu, G. (2009). ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2 (2), 438-469.
2. Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60 (2), 91-110.
3. J. Karhunen and J. Joutsensalo. Generalization of Principal Component Analysis, Optimization Problems and Neural Networks. *Neural Networks*, 8(4), 1995.
4. Ke, Y., Sukthankar, R. (2004). PCA-SIFT: A More Distinctive Representation for Local Image Descriptors. In *Computer Vision and Pattern Recognition (CVPR 2004)*, 27 June – 2 July 2004. IEEE, Vol. 2, 506-513.
5. Abdel-Hakim, A.E., Farag, A.A. (2006). CSIFT: A SIFT descriptor with Color Invariant Characteristics. In *Computer Vision and Pattern Recognition (CVPR 2006)*, 17-22 June 2006. IEEE, Vol. 2, 1978-1983.
6. Mortensen, E.N., Deng, H., Shapiro, L. (2005). A SIFT Descriptor with Global Context. In *Computer Vision and Pattern Recognition (CVPR 2005)*, 20-25 June 2005. IEEE, Vol. 1, 184-190.
7. Bay, H., Tuytelaars, T., Gool, L.V. (2006). SURF: Speeded Up Robust Features. In *Computer Vision – ECCV 2006 : 9th European Conference on Computer Vision*, 7-13 May 2006. Springer, Part II, 404-417.
8. Juan, L., Gwon, O. (2009). A comparison of SIFT, PCA-SIFT and SURF. *International Journal of Image Processing*, 3 (4), 143-152.
9. Mr. Amit Kr. Gautam, Ms. Twisha. "Improved Face Recognition Technique using SIFT", *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)*, 2014.
10. Shilpi Sonil, Raj Kumar Sahu2. "Face Recognition Based on Two Dimensional Principal Component Analysis (2DPCA) and Result Comparison with Different Classifiers", *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 2, Issue 10, 2013.